حسن زاده ، محمد؛ غفاری، سعید؛ زارعی، عاطفه ؛کمندی، حسین (۱۳۹۳). کارکرد عنوان و نشانی اینترنتی در بهبود ربط نتایج بازیابی اطلاعات. *پژوهش‌نامه کتابداری و اطلاع‌رسانی*، ۴ (۱)، ۲۴۰-۲۲۱.

کارکرد عنوان و نشانی اینترنتی در بهبود ربط نتایج بازیابی اطلاعات

دکترمحمد حسن زاده [۱]، دکتر سعید غفاری [۲]، عاطفه زارعی [۳]، حسین کمندی [۴]

تاریخ دریافت: ۹۱/۱۱/۲۱ تاریخ پذیرش: ۹۲/۱۲/۱۹

چکیده

هدف: *پژوهش حاضر به مقایسه میزان ربط نتایج حاصل از محدودسازی عبارت جست‌وجو به عنوان و نشانی اینترنتی توسط کارشناسان گروه‌های فنی مهندسی، علوم پایه و اقتصاد در استانداری همدان می‌پردازد. همچنین میزان حضور کلیدواژه‌های مورد جست‌وجو در عنوان و نشانی اینترنتی وب‌سایت‌های این گروه‌ها را مورد مقایسه قرار می‌دهد.*

روش: *جامعه آماری این پژوهش ۹۰ نفر از کارشناسان خبره استانداری همدان در سه گروه فنی مهندسی، علوم پایه و اقتصاد هستند که در حال انجام پروژه‌های تحقیقاتی و مطالعاتی برای توسعه استان بودند. آنان با استفاده از قسمت جست‌وجوی پیشرفته موتور کاوش گوگل، کلیدواژه‌های (یک و چند واژه‌ای) مورد نظر خود را، در قسمت نشانی اینترنتی و عنوان صفحات وب جست‌وجو کردند و میزان ربط نتایج بازیابی شده را در چک لیست‌هایی که به همین منظور تهیه شده بود، تعیین نمودند. برای دسته‌بندی و تحلیل داده‌ها از آمار توصیفی و استنباطی استفاده شد.*

یافته‌ها: *یافته‌های پژوهش نشان داد با افزایش تعداد کلیدواژه‌های مورد جست‌وجو، میزان هم‌خوانی عنوان و نشانی اینترنتی کاهش می‌یابد و هر چه هم‌خوانی بین عنوان و نشانی اینترنتی بیش‌تر باشد میزان ربط بالاتر می‌رود. بین تعداد کلیدواژه‌های موضوعی موجود در عنوان و نشانی اینترنتی گروه‌های مورد بررسی تفاوت معناداری وجود دارد. موتورهای کاوش با افزایش وزن‌دهی به کلیدواژه‌های موجود در نشانی اینترنتی می‌توانند نتایج مرتبط‌تری را رتبه‌بندی کرده و در اختیار کاربران قرار دهند.*

کلید واژها: *بازیابی اطلاعات، عنوان، کلیدواژه، موتورهای کاوش، میزان ربط، نشانی اینترنتی، صفحات وب*

---

۱. هیأت علمی گروه علوم اطلاعات و دانش شناسی دانشگاه تربیت مدرس تهران، hasanzadeh@modares.ac.ir

۲. عضو هیأت علمی علوم اطلاعات و دانش شناسی دانشگاه آزاد اسلامی واحد همدان، ghaffari130@ yahoo.com

۳. عضو هیأت علمی علوم اطلاعات و دانش شناسی دانشگاه آزاد اسلامی واحد همدان، Atefehzarei@gmail.com

۴. کارشناس ارشد علوم اطلاعات و دانش شناسی استانداری همدان، Kamandihossein@yahoo.com

مقدمه

بازیابی اطلاعات از شبکه جهان‌گستر وب و جزییات فرآیند اطلاع‌یابی در این محیط به دلایل متعددی مورد توجه پژوهشگران حوزه بازیابی اطلاعات قرار گرفته است. ویژگی‌های محیط وب از جمله یکپارچگی آن و امکان برقراری پیوند میان اطلاعات مختلف گرچه از امتیازات برجسته برای این محیط به شمار می‌آید در هنگام جست‌وجو و بازیابی در موتورهای کاوش تعداد نشانی‌های اینترنتی یا پیوندهایی که به منزله نتایج جست‌وجو ارائه می‌شوند معمولاً بسیار زیاد هستند. اما همه این نتایج مرتبط نیستند و حتی ممکن است عواملی مثل ابهام در زبان (مثل ماهیت خاص رسم‌الخط فارسی) باعث شود نتایج مناسبی به کاربر داده نشود. برای فراهم کردن دسترسی سریع و درعین‌حال به صفحات مناسب و این که صفحات با ارتباط موضوعی بیش‌تر، در الویت بالاتری قرار بگیرند، الگوریتم‌های جست‌وجو راهبردهای رتبه‌بندی مختلفی را به کار می‌برند. همچنین موتورهای کاوش از راهبردهای مختلفی برای کاهش زمان جست‌وجو و افزایش ربط نتایج استفاده می‌کنند. این نشانی اینترنتی و عنوان در صفحات وب است که کاربران را یاری می‌کنند کجا هستند و انتظار دارند که چه چیزی را پیدا کنند.

عنوان در هنگامی که کاربری سایتی را بازیابی می‌کند، استفاده می‌شود؛ بنابراین، عنوان مهم‌ترین جزء اطلاعات برای ذخیره و بازیابی در موتورهای کاوش است(Debolt, 2002). هنگام جست‌وجو در موتورهای کاوش صفحه‌های وب که واژه مورد جست‌وجو در عنوان آن‌ها به‌کاررفته باشد، معمولاً" به منزله مرتبط‌ترین مدارک نسبت به موضوع جست‌وجو در نظر گرفته می‌شود. همچنین نشانی‌های اینترنتی[۱] نیز در صفحات وب به منظور مکان‌یابی و دسترسی کاربران به اطلاعات در نظر گرفته‌شده‌اند. در اینترنت نیز برای دسترسی به واحدهای اطلاعات ، نیاز به یک نشانی است که ما را به آن واحد اطلاعاتی، راهنمائی کند و بسیاری از کاربران در هنگام تصمیم‌گیری در مورد استفاده از یک سایت به نشانی اینترنتی توجه می کنند. کاترل و گان[۲] در تحقیق خود که در نوامبر سال ۲۰۰۷ انجام دادند، بیان کردند که استفاده‌کنندگان ۲۴٪ از زمان جست‌وجوی خود را صرف نگاه کردن و تحلیل نشانی اینترنتی می‌کنند (در: Gardne، 2007) و به همین دلیل، با توجه به اهمیت عنوان و نشانی اینترنتی منابع وبی، بیشتر موتورهای کاوش امکان محدود کردن دامنه جست‌وجو به "عنوان" و" نشانی اینترنتی" را فراهم کرده‌اند.

---
[1] Uniform resource location (URL)
[2] Guan & Cutrell

بیان مسئله

در جست‌وجوی اطلاعات از طریق موتورهای کاوش تعداد بسیار زیادی نشانی اینترنتی بازیابی می‌شوند اما همه این نتایج مرتبط نیستند. این نتایج معمولاً چنان فراوان هستند که کاربران با مرور چند صفحه اول از مرور صفحات دیگر منصرف شده و مجبور می‌شوند به رتبه‌بندی موتور کاوش اعتماد کنند. برای فراهم کردن دسترسی سریع به صفحات مرتبط و اینکه صفحات با ارتباط موضوعی بیشتر در اولویت بالاتری قرار بگیرند، الگوریتم‌های جست‌وجو، راهبردهای رتبه‌بندی مختلفی را به کار می‌برند. همین طور موتورهای کاوش از راه‌های مختلفی برای کاهش زمان جست‌وجو و افزایش ربط نتایج استفاده می‌کنند و این موتورها امکانات گوناگونی را برای محدودسازی دامنه جست‌وجو و در نتیجه بالا بردن احتمال ربط نتایج بازیابی در نظر می‌گیرند. یکی از این امکانات، محدود کردن کلیدواژه‌ها و عبارات جست‌وجو به عنوان و یا نشانی اینترنتی است که تعداد نتایج جست‌وجو را به طور معناداری کاهش داده و درعین‌حال احتمال ربط آنان را افزایش می‌دهد (Fattahi, 2006). موتورهای کاوش مدعی هستند که آنچه که از طریق محدود کردن عنوان و نشانی بازیابی می‌شود، از ربط بیش‌تری برخوردارند (Greg, 2004). در این پژوهش سعی شده است تا مشخص شود که تا چه میزان بین دیدگاه کاربران مورد پژوهش با دیدگاه نظام بازیابی اطلاعات در وب در مورد "ربط نتایج" همسانی وجود دارد و آیا این نتایج از ربط مناسبی از دیدگاه کاربر نیز برخوردار هستند؟ کاربران گروه‌های موضوعی مختلف با جست‌وجو در کدام قسمت(جست‌وجو در عنوان و یا از طریق نشانی اینترنتی) به نتایج مرتبط تری دست می‌یابند؟ میزان اشتراک نتایج بازیابی شده از طریق عنوان و نشانی اینترنتی چگونه است ؟

اهداف پژوهش

هدف کلی پژوهش: هدف پژوهش حاضر تطبیق و مقایسه جست‌وجو در "عنوان" و "نشانی اینترنتی" در بهبود ربط نتایج بازیابی شده اطلاعات وبی توسط کارشناسان گروه‌های مختلف موضوعی مورد پژوهش می‌باشد.

اهداف فرعی پژوهش

۱. تعیین نتایج مرتبط توسط کاربران گروه‌های موضوعی مورد پژوهش با جست‌وجو در "عنوان" و «نشانی اینترنتی»

۲. شناسایی تفاوت‌ها در میزان ربط نتایج جست‌وجوی مبتنی بر "عنوان" و "نشانی اینترنتی" در گروه‌های موضوعی مورد پژوهش

۳. مقایسه تفاوت دیدگاه‌های کاربران گروه‌های موضوعی با نظام بازیابی اطلاعات در مورد "ربط نتایج"

سؤالات پژوهش

۱- ربط نتایج جست‌وجوی مبتنی بر "عنوان " و مبتنی بر " نشانی اینترنتی" منابع وبی، در گروه‌های موضوعی مختلف به چه میزان است ؟

۲- آیا میزان ربط نتایج جست‌وجو بر اساس «عنوان» بین کاربران گروه‌های موضوعی مورد پژوهش به صورت معنی‌داری متفاوت است؟

۳- آیا میزان ربط نتایج جست‌وجو بر اساس «نشانی» بین کاربران گروه‌های موضوعی مورد پژوهش به صورت معنی‌داری متفاوت است؟

۴- آیا تفاوتی بین دیدگاه کاربران موضوعی مورد پژوهش با دیدگاه نظام بازیابی اطلاعات وبی وجود دارد؟

فرضیه‌های پژوهش

۱- میزان هم‌خوانی کلیدواژه‌های نشانی اینترنتی و عنوان با افزایش تعداد کلیدواژه‌های مورد جست‌وجو به صورت معنی‌داری متفاوت است.

۲- بین گروه‌های موضوعی مورد پژوهش به لحاظ تعداد کلیدواژه‌های موضوعی موجود در عنوان و نشانی اینترنتی تفاوت وجود دارد.

پیشینه پژوهش

در مورد ربط و بازیابی اطلاعات در موتورهای کاوش پژوهش‌های زیادی انجام گرفته است اما کارکرد عنوان و نشانی اینترنتی در بهبود ربط نتایج بازیابی اطلاعات که بحث اصلی این پژوهش است کم‌تر توجه قرار گرفته است. کمیته سئو در سال ۲۰۰۴ تحقیقی بر روی تأثیرگذاری کلیدواژه‌های موجود در نشانی اینترنتی در دو موتور کاوش یاهو و گوگل انجام داد و مشخص شد که هر دو موتور، حضور کلیدواژه‌ها در نشانی اینترنتی را عاملی در رتبه‌بندی در نظر می‌گیرند(SEO Committee, 2004). کیم

وکنگ نیز درهمان سال تحقیقی بر روی بازیابی اطلاعات در رابطه با پرسش‌های کاربر انجام دادند. آن‌ها یک روش رده‌بندی برای پرسش‌های کاربر ارائه دادند و بعد از آن‌که توانستند پرسش‌های کاربر را رده بندی کنند، از الگوریتم متفاوت برای ارائه نتایج مرتبط‌تر استفاده کردند. برای این منظور آن‌ها بر سه مورد تأکید کردند: اطلاعات متن، اطلاعات پیوندها و اطلاعات نشانی‌های اینترنتی. و بیان کردند که اطلاعات موجود نشانی‌های اینترنتی و پیوندها برای یافتن صفحات خانگی مرتبط‌تر مفید هستند (Kang, In-Ho, & Gil Chang Kim, 2004). ویدمن و استرومفر تأثیر تعداد کلیدواژه‌های به‌کاررفته توسط کاربران در جست‌وجو و میزان موفقیت بازیابی را بررسی کردند. تحلیل آنان نشان می‌دهد که تفاوت زیادی در میزان موفقیت بین کاربرانی که فقط از یک کلیدواژه استفاده می‌کنند (میزان موفقیت کم) و کسانی که دو یا بیشتر کلیدواژه بکار می‌برند، وجود دارد. به طور تخصصی‌تر، به دست آوردن نتیجه قابل‌قبول در جست‌وجو، زمانی که از دو یا چند کلیدواژه استفاده می‌شود، ۵/۵ برابر زمانی است که یک کلیدواژه مورد جست‌وجو قرار می‌گیرد. (Weideman & Strümpfer, 2004) . شین و کارگر در سال ۲۰۰۴ باهدف تعیین الگوریتمی برای رده‌بندی وب مشخص کردند که اگر به جای محتوای متن، به نشانی اینترنتی پیوندها و مکان بصری پیوندها توجه شود، رده‌بندی وب امکان‌پذیر می‌شود. آن‌ها توانستند الگوریتمی برای رده‌بندی وب با توجه به نشانی اینترنتی و محل قرارگیری پیوندها در منابع وبی، پیشنهاد دهند(Shin & Karger, 2004).

مشابه پژوهشی که شین و کارگر انجام دادند؛ کاتو، کروهاشی و کنتارودر سال ۲۰۰۸ نیز پژوهشی درباره رده‌بندی منابع وب با هدف توسعه یک روش رده‌بندی جدید انجام دادند در این پژوهش برای آزمایش روشی در رده‌بندی کردن مدارک وب از عنوان و نشانی اینترنتی استفاده کردند و نتایج حاصله نشان داد که روش پیشنهادی با استفاده از عوامل یادشده از بیش از ۵۰٪ دقت برخوردار است (Yoshikiyo & Kurohashi & kentaro), 2008). فتاحی،کول فلیچر و ویلسون در سال ۲۰۰۸ پژوهشی درباره عنوان و نشانی اینترنتی با هدف بسط جست‌وجو انجام دادند و تأثیر کلیدواژه‌های عمومی را در عنوان و نشانی اینترنتی در بازیابی اطلاعات مرتبط در یک موتور کاوش (گوگل) نشان دادند. نتایج پژوهش آن‌ها که بسط جست‌وجو در عنوان و در نشانی اینترنتی، با استفاده از واژه‌های عمومی که همراه کلیدواژه‌های موضوعی در عنوان و نشانی اینترنتی منابع و متون موجود در وب ظاهر می‌شوند، می‌تواند موجب افزایش میزان دقت و ربط در نتایج بازیابی در موتورهای کاوش شود(Fattahi Concepción S Wilson & Fletcher Cole, 2008).

در ایران توسط برخی پژوهشگران علوم اطلاع‌رسانی، پژوهش‌های مربوط به موتورهای کاوش انجام گرفته است که از جمله آن‌ها می‌توان به پژوهش‌های علیجانی و طیبی عراقی به ترتیب در سال‌های ۱۳۸۱ و۱۳۸۲ اشاره کرد (Allijani,2004) و (Taiebi aragi, 2005). برخی از آنان نیز معیارهای دقت و بازیافت را برای مقایسه موتورهای کاوش در نظر گرفته‌اند و درعین‌حال در مورد کارکرد عنوان و نشانی اینترنتی پژوهش‌های بسیار اندکی صورت گرفته است.

فتاحی سال ۱۳۸۵ در تحقیق خود درباره عنوان و نشانی اینترنتی که با هدف بسط جست‌وجو انجام گرفته بود، تأثیر کلیدواژه‌های غیر موضوعی (عمومی) در عنوان و نشانی اینترنتی را در بازیابی اطلاعات مرتبط در یک موتور کاوش (گوگل) نشان داد (Fattahi, 2006). علامی و فتاحی در مقاله خود با عنوان تأثیر عنوان و نشانی اینترنتی در میزان ربط نتایج حاصل از بازیابی اطلاعات در موتورهای کاوش در دو حوزه علوم کشاورزی و علوم انسانی (در کلید واژه‌های زبان انگلیسی) بیان کردند که جستجو بر اساس نشانی اینترنتی در حوزه‌های موضوعی علوم کشاورزی نتایج مرتبط‌تری را به دنبال دارد و تفاوت معنی‌داری در بین تعداد کلیدواژه‌های موجود در عنوان و نشانی اینترنتی دو حوزه موضوعی وجود ندارد(Allami & Fattahi, 2012).

به طور کلی یافته‌های این پژوهش به کاربران درگروه‌های موضوعی اقتصاد، علوم پایه و فنی مهندسی در زمینه جست‌وجوی منابع مرتبط‌تر و مفیدتر وبی در ابزارهای کاوش با واسط فارسی و نیز در زمینه طراحی بهتر و مناسب‌تر صفحات وب به طراحان وب‌های فارسی‌زبان با توجه به میزان حضور کلیدواژه در مکان‌های مختلف، یاری می‌رساند و می‌تواند طراحان موتورهای کاوش را در تدوین الگوریتم‌های جست‌وجو برای موتورهای کاوش و نیز صفحه رابط جست‌وجو که به شکل پیش‌فرض در اختیار کاربر قرار داده می‌شود یاری رساند و همچنین به شیوه‌های بهتر نمایه‌سازی موتورهای کاوش به منظور بازیابی مؤثر اطلاعات و نیز هدایت کاربر در بازیابی اطلاعات مرتبط‌تر، در کلیدواژه‌های موضوعی فارسی‌زبان در جست‌وجو به روش عنوان و نشانی اینترنتی کمک کند و تفاوت این پژوهش با سایر پژوهش‌های انجام‌شده در استفاده از زبان کلیدواژه‌های مورد جست‌وجو (که با واسط فارسی) در این پژوهش انجام گرفته است و نتیجه آن با نتایج پژوهش‌هایی که از کلیدواژه‌های زبان انگلیسی استفاده‌شده مورد مقایسه قرار گرفته است.

روش‌شناسی پژوهش

این پژوهش از نوع کاربردی است و با روش پیمایشی به بررسی میزان هم‌خوانی و عملکرد تعداد کلیدواژه‌ها در میزان هم‌خوانی و ناهم‌خوانی و همچنین به بررسی وب سایت‌ها از دیدگاه کاربران می‌پردازد. جامعه آماری پژوهش حاضر ۹۰ نفر از کارشناسان گروه‌های مورد پژوهش، که از این بین ۳۰ نفر به گروه فنی مهندسی ، ۳۰ نفر گروه علوم پایه و۳۰ نفر از گروه اقتصاد در استانداری همدان که آشنایی کاملی با وب و به زبان انگلیسی نیز مسلط بودند برای بررسی میزان ربط نتایج جست‌وجو شده انتخاب شدند.

جامعه دیگر این پژوهش وب سایت‌های حاصل از جست‌وجو در موتور کاوش توسط کارشناسان گروه‌های مختلف می‌باشد که توسط کاربران مورد بررسی و تحلیل قرار گرفتند و برای گردآوری داده‌های مورد نیاز این پژوهش از چک لیست‌ها استفاده شد و همچنین به منظور تعیین موارد همپوشان در نتایج جست‌وجوهای مبتنی بر عنوان و نشانی اینترنتی و بررسی رتبه آنان در هر دو شیوه جست‌وجو چک لیست‌هایی طراحی شد و موارد همپوشان در جست‌وجوی مبتنی بر عنوان و نشانی توسط پژوهشگر تکمیل و ثبت شد. برای تحلیل داده‌ها در این پژوهش از روش آمار توصیفی (فراوانی، میانگین و انحراف استاندارد) و آمار استنباطی (آزمون مقایسه میانگین سه جامعه، آزمون همبستگی پیرسون، آزمون تک متغیری کالموگورف اسمیرنف و آزمون فرض یک طرفه) استفاده شد.

در تحلیل داده‌ها از آنجا که داده‌ها رتبه‌ای بودند، نمره‌های ۰ تا ۴ برای آن‌ها در نظر گرفته شد ابتدا از هر یک از آزمودنی‌ها خواسته شد تا هر کدام پنج کلیدواژه موضوعی فارسی (دارای یک تا چند واژه) را بر اساس آخرین نیاز اطلاعاتی خود در نظر بگیرند و ابتدا در گزینه مربوط به "عنوان" در منوی (صفحه) جست‌وجوی پیشرفته، جست‌وجو را انجام دادند. از میان اقلام بازیابی شده تعداد ده سایت اول را مورد بررسی قرار داده و تعداد موارد بازیابی شده و تعداد موارد مرتبط و میزان ارتباط ده مورد اول توسط آزمودنی‌ها در چک لیست ثبت شد، سپس میانگین رتبه‌های تعیین‌شده برای هر کلیدواژه، توسط پژوهشگر محاسبه و از کاربران خواسته شد جست‌وجوی خود را عیناً با همان کلیدواژه‌های موضوعی در گزینه مربوط به "نشانی اینترنتی"تکرار کنند.

تعداد موارد بازیابی شده و تعداد موارد مرتبط و میزان ارتباط مدارک مانند مرحله قبل ثبت شد. برای هر موضوع ده نتیجه برای جست‌وجو در "عنوان" و ده نتیجه برای جست‌وجوی در "نشانی اینترنتی" مورد بررسی قرار گرفت که در مجموع بیست صفحه وب برای هر موضوع مورد جست‌وجو واقع شد و هر کاربر در حدود صد صفحه وب را مورد بررسی قرارداد. این پژوهش به بررسی تعداد کلیدواژه‌های

موضوعی موجود در عنوان و در نشانی اینترنتی میان گروه‌های موضوعی مورد پژوهش نیز پرداخت. برای بررسی تعداد کلیدواژه‌های موضوعی تعداد سی کلیدواژه از هر گروه انتخاب شد. سپس در سی نتیجه بازیابی شده اول به شمارش کلیدواژه‌های موضوعی آن‌ها پرداخته شد. آمار نشان می‌دهد بسیاری از کاربران دریافته‌اند که در اغلب موارد صفحه نخست نتایج جست‌وجو می‌تواند خواسته آن‌ها را برآورده کند (Web wersad, 2006).

بنابراین تعداد کلیدواژه‌ها و همچنین میزان هم‌خوانی در ده نتیجه اول بررسی شدند و همچنین، این پژوهش به بررسی میزان هم‌خوانی کلیدواژه‌های نشانی اینترنتی منابع وبی با کلیدواژه‌های عنوان آن‌ها بوده است. برای این منظور ابتدا کلیدواژه‌های مورد جست‌وجو در گروه‌های مورد پژوهش به دو قسمت یک واژه‌ای و چند واژه‌ای تقسیم شد و سپس از هر کدام از گروه‌ها تعداد سی کلیدواژه انتخاب شدند.

بنابراین در دو قسمت یک واژه‌ای و چند واژه‌ای ۶۰ کلیدواژه مورد بررسی قرار گرفت و میزان هم‌خوانی کلیدواژه‌های موجود در نشانی اینترنتی منابع وبی با کلیدواژه‌های عنوان آن‌ها به طور جداگانه ثبت شد و سپس میانگین رتبه‌های تعیین‌شده برای هر کلیدواژه مانند مورد قبل توسط پژوهشگر محاسبه و میانگین رتبه‌ها وارد نرم‌افزار SPSS شد و برای تعیین روایی و پایایی سیاهه‌های ارزیابی که توسط پژوهشگر تدوین شده است، از روش اعتبار محتوا و مشورت با صاحب‌نظران استفاده شد. بنابراین در تنظیم سیاهه‌های وارسی برای تصحیح مشکلات و پیچیدگی‌های آن از نظر برخی صاحبنظران استفاده شد و برای سنجش پایایی ابزار گردآوری‌شده ضریب آلفای کرونباخ با استفاده از SPSS گرفته شد. ضریب آلفای کرونباخ برای سیاهه مورد نظر ۰/۸۲۹ می‌باشد که بیانگر پایایی ابزار مورد استفاده می‌باشد.

یافته‌های پژوهش:

سؤال۱- میزان ربط نتایج جست‌وجوی مبتنی بر "عنوان" و مبتنی بر " نشانی اینترنتی" منابع وبی، در گروه‌های موضوعی مختلف به چه صورت است ؟

پیش از تعیین میزان ربط، ابتدا با استفاده از آزمون تک متغیری کالموگورف اسمیرنف برای تعیین میزان نرمال بودن داده‌های عنوان و نشانی اینترنتی در گروه‌های موضوعی استفاده شد. با توجه به مقادیر p به دست آمده از آزمون فوق در تمام گروه‌ها که از مقدار ۰/۰۵ بزرگ‌تر است نرمال بودن توزیع داده‌ها را نشان می‌دهد، به همین دلیل، برای تعیین میزان ربط نتایج حاصل از جست‌وجو در دو مقوله عنوان و نشانی اینترنتی از آزمون t استفاده شد.

جدول۱.میزان ربط نتایج حاصل از جست وجو در عنوان و نشانی اینترنتی با استفاده از آزمون t زوجی

گروه‌ها	میانگین اختلاف	انحراف معیار اختلاف	مقدار آما ره	p-value
اقتصاد	۰/۴۵۵	۰/۱۳۸	۱۰/۸۰	۰/۰۰۰
علوم پایه	۰/۲۴۲	۰/۲۸۰	۲/۸۶۵	۰/۰۱۷
فنی و مهندسی	۰/۹۰۷	۰/۲۱۴	۱۴/۰۶	۰/۰۰۰

در جدول(۱) نتایج به دست آمده از آزمون t زوجی در تمام گروه‌ها از مقدار ۰/۰۵ کوچک‌تر است، نشان‌دهنده آن است که در گروه‌های بالا بین کارکرد عنوان و نشانی اینترنتی در نتایج بازیابی اطلاعات مرتبط تفاوت معنی‌داری وجود دارد.

سؤال۲- آیا میزان ربط نتایج جست‌وجو بر اساس «عنوان» بین کاربران گروه‌های موضوعی مورد پژوهش به صورت معناداری متفاوت است؟

میزان ربط نتایج حاصل از جست‌وجو بر اساس عنوان، با استفاده از داده‌های به دست آمده از سیاهه وارسی شماره یک تعیین شد. سپس برای بررسی معنی تفاوت ربط در بین گروه‌های مختلف از آزمون تحلیل واریانس استفاده گردید.

جدول۲. تحلیل واریانس برای تفاوت میزان ربط نتایج حاصل از جست و جوی گروه‌های مورد مطالعه در "عنوان"

منبع تغییر	مجموع مربعات خطا	درجه آزادی	میانگین مربعات خطا	مقدار آماره F	مقدارP-valu
بین گروه‌ها	۱/۵۸	۲	۰/۹۲۴	۳۹/۴۳۶	۰/۰۰۰
خطا	۰/۷۰۴	۳۰	۰/۰۲۳		
کل	۲/۵۵۴	۳۲			

در جدول (۲) مقدار P در آزمون تحلیل واریانس یک طرفه برابر با ۰/۰۰۰ می‌باشد و با توجه به اینکه این مقدار از ۰/۰۵ کوچک‌تر است، نشان‌دهنده آن است که در میان گروه‌های مورد پژوهش از دید کاربران

تفاوت معناداری به لحاظ نتایج حاصل از جست‌وجو در عنوان مشاهده می‌شود و سپس با استفاده از آزمون تعقیبی دانکن الگوی تفاوت مشخص شد (جدول ۳).

جدول۳. آزمون تعقیبی دانکن برای نتایج حاصل از جست‌وجو در "عنوان"

گروه‌های مورد پژوهش	N	مقدار آلفا برای زیرمجموعه‌ها = ۰/۰۵	
		۱	۲
فنی مهندسی	۱۱	۲/۸۷۰۰	
اقتصاد	۱۱		۳/۳۶۹۱
علوم پایه	۱۱		۳/۳۷۵۵
سطح معناداری		۱/۰۰۰	۰/۹۲۳

در جدول (۳) با استفاده از آزمون تعقیبی دانکن برای نتایج حاصل از جست‌وجو در "عنوان سطح معنی‌داری از دید کاربران گروه فنی مهندسی ربط نتایج حاصل از جست‌وجو در" عنوان "کمتر از گروه‌های اقتصاد و علوم پایه است که این تفاوت، بین گروه " فنی مهندسی" با سایر گروه‌هاست و بین میانگین‌های دو گروه اقتصاد و علوم پایه تفاوت از لحاظ آماری معنی‌دار نیست.

سؤال ۳ – آیا میزان ربط نتایج جست‌وجو بر اساس «نشانی» بین کاربران گروه‌های موضوعی مورد پژوهش به صورت معناداری متفاوت است؟

میزان ربط نتایج حاصل از جست‌وجو بر اساس نشانی اینترنتی، با استفاده از داده‌های به دست آمده از سیاهه وارسی شماره یک تعیین شد. سپس برای بررسی معنی تفاوت ربط در بین گروه‌های مختلف از آزمون تحلیل واریانس استفاده گردید.

جدول۴. جدول آنوا " برای نتایج حاصل از جست‌وجو در " نشانی اینترنتی "

منبع تغییر	مجموع مربعات خطا	درجه آزادی	میانگین مربعات خطا	مقدار آماره F	P-value
بین گروه‌ها	۴/۶۴۴	۲	۲/۳۲۲	۶۶/۰۴۹	۰/۰۰۰
خطا	۱/۰۵۵	۳۰	۰/۰۳۵		
کل	۵/۶۹۹	۳۲			

در جدول (۴) مقدار P در آزمون تحلیل واریانس یک طرفه برابر با ۰/۰۰۰ می‌باشد و با توجه به اینکه این مقدار از ۰/۰۵ کوچک‌تر است، نشان‌دهنده آن است که در میان گروه‌های مورد پژوهش از دید کاربران تفاوت معناداری به لحاظ نتایج حاصل از جست‌وجو در نشانی اینترنتی مشاهده می‌شود. سپس با استفاده از آزمون تعقیبی دانکن الگوی تفاوت مشخص شد (جدول ۵).

جدول۵.آزمون تعقیبی دانکن برای نتایج حاصل از جست‌وجو در " نشانی اینترنتی"

گروه‌ها	N	مقدار آلفا برای زیرمجموعه‌ها = ۰/۰۵	
		۱	۲
اقتصاد	۱۱	۲/۹۱۳۶	
علوم پایه	۱۱		۳/۶۱۷۳
فنی مهندسی	۱۱		۳/۷۷۳۷
سطح معنی‌داری		۰/۰۰۰	۰/۰۵۴

در جدول(۵) با استفاده از آزمون تعقیبی دانکن برای نتایج حاصل از جست‌وجو در " نشانی اینترنتی" سطح معنی‌داری در دو گروه فنی مهندسی و علوم پایه برابر با ۰/۰۵۴ می‌باشد و از دید کاربران گروه‌های فوق ربط نتایج حاصل از جست‌وجو در" عنوان " کمتر از گروه‌های اقتصاد است و این تفاوت، بین گروه " اقتصاد " با سایر گروه‌هاست و بین میانگین‌های دو گروه فنی مهندسی و علوم پایه تفاوت از لحاظ آماری معنی‌دار نیست.

سؤال ۴- آیا تفاوتی بین دیدگاه کاربران موضوعی مورد پژوهش با دیدگاه نظام بازیابی اطلاعات وبی وجود دارد؟

جدول۶ .میزان ربط نتایج حاصل از جست‌وجو در عنوان و نشانی اینترنتی از دید کاربر و نظام

گروه‌ها	نوع جست‌وجو	دید کاربر					نظام	میانگین	انحراف استاندارد
		بسیار مرتبط	مرتبط	متوسط	کم ربط	غیر مرتبط			
فنی مهندسی	عنوان	۵۲	۳۶	۳۶	۶۸	۵۲		۲/۸۷	۲/۱۱
	نشانی	۷۷	۸۷	۴۱	۲۹	۱۱	۲۴۵	۳,۷۸	۱,۳۱
اقتصاد	عنوان	۸۱	۶۷	۳۹	۸۱	۲۱	۲۸۹	۳/۳۷	۱,۷۹
	نشانی	۳۶	۴۸	۶	۱۱۲	۱۶	۲۷۷	۲،۹۱	۱,۳۲
پایه	عنوان	۴۰	۸۹	۴۳	۵۴	۱۳	۲۳۹	۳,۳۷	۱,۳۵
	نشانی	۴۶	۹۰	۴۱	۲۰	۱۷	۲۱۴	۳,۶۱	۱,۳۱

در جداول (۶) به مقایسه میزان ربط از دید کاربر و نظام پرداخته شده است. نتایج مرتبط بازیابی شده از دید نظام و میزان ربط همان نتایج از دید کاربر را نشان می‌دهند.

در این جدول موارد یک تا ده نتایج رتبه‌بندی شده توسط موتور کاوش هستند که بر اساس الگوریتم رتبه‌بندی مرتبط‌ترین موارد توسط نظام تشخیص داده‌شده‌اند و همان نتایج از دید کاربر مورد سنجش قرارگرفته و نتایج حاصل از جست‌وجو در عنوان و نشانی اینترنتی در گروه‌های مورد پژوهش از دید کاربر و نظام آورده شده است و در گروه فنی و مهندسی از ۲۴۴ مورد که از دید نظام مرتبط‌ترین گزینه در جست‌وجوی مبتنی بر عنوان هستند، ۱۲۴ مورد (۰/۵۰) نیز از دید کاربر مرتبط تلقی شده‌اند و نیز در جست‌وجوی مبتنی بر نشانی اینترنتی از ۲۴۵ مورد که از دید نظام به عنوان مرتبط‌ترین گزینه انتخاب‌شده‌اند، ۲۰۵ مورد (۸۳٪) نیز از سوی کاربر مرتبط تلقی شده‌اند که این امر خود نیز بیانگر افزایش ربط در جست‌وجوی مبتنی بر نشانی اینترنتی نسبت به عنوان در گروه موضوعی خاص مانند گروه فنی مهندسی می‌باشد و در گروه اقتصاد از ۲۸۹ مورد که از دید نظام مرتبط‌ترین گزینه در جست‌وجوی مبتنی

بر عنوان هستند، ۱۸۷ مورد (۶۵٪) نیز از دید کاربر مرتبط تلقی شده‌اند و نیز در جست‌وجوی مبتنی بر نشانی اینترنتی از ۲۷۷ مورد که از دید نظام به عنوان مرتبط‌ترین گزینه انتخاب‌شده‌اند، ۱۴۹ مورد (۵۳٪) از سوی کاربر مرتبط تلقی شده‌اند و در گروه علوم پایه از ۲۳۹ مورد که از دید نظام مرتبط‌ترین گزینه در جست‌وجوی مبتنی بر عنوان هستند، ۱۷۲ مورد (۷۱٪) نیز از دید کاربر مرتبط تلقی شده‌اند و نیز در جست‌وجوی مبتنی بر نشانی اینترنتی از ۲۱۴ مورد که از دید نظام به عنوان مرتبط‌ترین گزینه انتخاب‌شده‌اند، ۱۷۷ مورد (۸۲٪) نیز از سوی کاربر مرتبط تلقی شده‌اند و سپس برای مقایسه بررسی تفاوت میزان ربط نتایج حاصل از دیدگاه کاربران موضوعی مختلف با دیدگاه نظام بازیابی اطلاعات وبی از آزمون t استفاده می‌شود.

جدول۷ آزمون t برای مقایسه تفاوت میزان ربط در گروه‌ها از دید کاربران و نظام بازیابی اطلاعات وبی

گروه	نوع جست‌وجو	میانگین	انحراف معیار	مقدار آماره	p-value
فنی مهندسی	عنوان	۸۷/۲	۴۵۷/۱	−۲۲/۸۴۹	۰/۰۰۰
	نشانی	۸۰/۳	۱۷۶/۱	−۱۵/۵۰۸	۰/۰۰۰
اقتصاد	عنوان	۳۷/۳	۳۴۰/۱	−۲۰/۷۲۰	۰/۰۰۰
	نشانی	۹۱/۲	۱۵۱/۱	−۳۰/۱۶۱	۰/۰۰۰
علوم پایه	عنوان	۳۷/۳	۱۶۳/۱	−۲۱/۶۳۸	۰/۰۰۰
	نشانی	۴۶/۳	۱۹۵/۱	−۱۹/۸۱۲	۰/۰۰۰

در جدول(۷) مقدار p در آزمون t در تمام گروه‌ها در هر دو نوع جست‌وجو سطح معنی‌داری، برابر ۰/۰۰۰ است که از سطح خطای ما یعنی ۰/۰۵ کمتر است، نشان‌دهنده آن است که میان گروه‌های مورد پژوهش از دید کاربر و نظام، تفاوت معناداری در بین نتایج حاصل از جست‌وجو(عنوانی و نشانی اینترنتی) وجود دارد.

آزمون فرضیه‌های پژوهش

فرضیه شماره۱: میزان هم‌خوانی کلیدواژه‌های نشانی اینترنتی و عنوان با افزایش تعداد کلیدواژه‌های مورد جست‌وجو به صورت معنی‌داری متفاوت است؟

ابتدا سی کلیدواژه یک واژه‌ای و سی کلیدواژه چند واژه‌ای در مجموع شصت کلیدواژه از گروه‌های مورد پژوهش انتخاب و در مورد هر کلیدواژه ده صفحه وب مورد بررسی قرار گرفت، سپس برای بررسی تفاوت میزان هم‌خوانی و عدم هم‌خوانی کلیدواژه‌های یک و چند واژه‌ای با استفاده از آزمون استقلال خی دو بین گروه‌های مختلف استفاده گردید.

جدول۸. آزمون استقلال خی دو میزان هم خوانی و عدم همخوانی کلیدواژه‌های یک و چند واژه‌ای

	مقدار	درجه آزادی	سطح معنی‌داری متقارن	سطح معنی‌داری دقیق دو طرفه	سطح معنی‌داری دقیق یک طرفه
کای دو دقیق	۵۳۸/۱۴	۱	۰/۰۰۰		
تصحیح پیوستگی	۸۰۲/۱۳	۱	۰/۰۰۰		
نسبت درست نمایی	۷۰۴/۱۴	۱	۰/۰۰۰		
آزمون دقیق فشر تست				۰/۰۰۰	۰/۰۰۰
ارتباط خطی	۵۱۳/۱۴	۱	۰/۰۰۰		
تعداد	۶۰۰				

مطابق با جدول (۸)، در آزمون استقلال خی دو با توجه به نتایج به دست آمده سطح معنی‌داری برابر ۰/۰۰۰ است که از سطح خطای ۰/۰۵ کمتر است. به این معنا که هر کدام از آماره‌های خی دو دقیق، تصحیح پیوستگی و موارد را که در نظر بگیریم، سطح معنی‌داری برابر با ۰/۰۰۰ است. این امر نشان می‌دهد که فرض مستقل بودن نتایج حاصل از" یک واژه‌ای یا چند واژه‌ای بودن کلیدواژه‌ها "از "هم‌خوانی یا عدم هم خوانی" رد می‌شود، سپس برای بررسی میزان هم‌خوانی کلیدواژه‌های یک و چند واژه‌ای با استفاده از آزمون ضرایب همبستگی پیرسون و اسپیرمن الگوی تفاوت مشخص شد (جدول ۹).

جدول۹ . ضرایب همبستگی اسپیرمن برای میزان هم خوانی کلیدواژه‌های یک و چند واژه‌ای

	مقدار	انحراف استاندارد	Tتقریبی	p-value
همبستگی اسپیرمن	-/۱۵۶	/۰۴۰	۳/۸۵۳	c۰/۰۰۰

تعداد	۶۰۰			

در (جدول۹) داده‌های به دست آمده در هر دو آزمون ضریب همبستگی پیر سون و اسپیرمن سطح معنی‌داری تقریبی کوچک‌تر از ۰/۰۵ می‌باشد و مقدار سطح معنی‌داری تقریبی = ۰/۰۰۰ c می‌باشد؛ و با توجه به منفی بودن ضرایب همبستگی آزمون‌ها نشان‌دهنده آن است که با افزایش تعداد کلیدواژه‌ها، میزان هم‌خوانی کاهش می‌یابد.

فرضیه شماره ۲ : بین گروه‌های موضوعی مورد پژوهش به لحاظ تعداد کلیدواژه‌های موضوعی موجود در عنوان و نشانی اینترنتی تفاوت وجود دارد.

ابتدا از آزمون تک متغیری کالموگورف اسمیرنف برای تعیین میزان نرمال بودن تعداد کلیدواژه‌های موضوعی موجود در داده‌های عنوان و نشانی اینترنتی در گروه‌های موضوعی استفاده شد. با توجه به مقادیر p به دست آمده از آزمون فوق در تمام گروه‌ها که از مقدار ۰/۰۵ کوچک‌تر است نرمال نبودن توزیع داده‌ها ثابت شد. نتیجه می گیریم که جامعه آماری نرمال نیست و به همین دلیل برای مقایسه میانگین چند جامعه از آزمون کروکسال- والیس استفاده می‌شود.

جدول۱۰ . آزمون کروکسال- والیس برای بررسی تعداد کلیدواژه‌های موضوعی موجود در جست‌وجوی عنوان توسط هر یک از گروه‌های مورد پژوهش

گروه موضوعی	تعداد کلیدواژه	میانگین رتبه	مقدار آ ما ره	p-value
اقتصاد	۳۰	۵۷/۹۵	۱۰/۲۶۸	۰/۰۰۶
علوم پایه	۳۰	۴۰/۰۰		
فنی مهندسی	۳۰	۳۸/۵۵		

درجدول (۱۰) میانگین و مقدار آماره تعداد کلیدواژه‌های موضوعی موجود در عنوان گروه‌های موضوعی را نشان می‌دهد که مقدار p برابر ۰/۰۰۶ می‌باشد،که گویای آن است در بین گروه‌های مورد پژوهش به لحاظ تعداد کلیدواژه‌های موضوعی موجود در عنوان تفاوت معنادار وجود دارد و در ادامه، این تفاوت در جست‌وجوی نشانی اینترنتی نیز با استفاده از آزمون کروکسال- والیس بررسی شد (جدول۱۱).

جدول۱۱. آزمون کروکسال- والیس برای بررسی تعداد کلیدواژه‌های موضوعی موجود در نشانی اینترنتی توسط هر یک از گروه‌های مورد پژوهش

گروه موضوعی	تعداد کلیدواژه	میانگین رتبه	مقدار آماره	p-value
اقتصاد	۳۰	۵۸/۶۷	۱۱/۵۲۵	۰/۰۰۳
علوم پایه	۳۰	۳۹/۹۲		
فنی مهندسی	۳۰	۳۷/۹۲		

در (جدول ۱۱) میانگین و مقدار آماره تعداد کلیدواژه‌های موضوعی موجود در نشانی اینترنتی گروه‌های موضوعی مورد بررسی را نشان می‌دهد. با توجه به اینکه مقدار p ۰/۰۰۳ می‌باشد و این مقدار از ۰/۰۵ کوچک‌تر است، فرض صفر(برابری میانگین‌ها) رد می‌شود و میان گروه‌های مورد پژوهش به لحاظ تعداد کلیدواژه‌های موضوعی موجود در نشانی اینترنتی تفاوت معنی‌دار وجود دارد.

نتیجه

هر چند که ربط امری نسبی است، ولی در نظر گرفتن عوامل مختلف تأثیرگذار بر بهبود ربط، می‌تواند تا حدودی مسئله ربط ظاهری یا تطابق میان واژگان را حل کند و تعداد کلیدواژه مورد جست‌وجو می‌تواند بر بهبود ربط نتایج جست‌وجو اثرگذار باشد. بر این اساس، بسیاری از موتورهای کاوش ، گزینه‌هایی برای محدود کردن واژه‌ها و عبارات مورد جست‌وجوی کاربران در مکان‌های خاص (مانند عنوان، متن، نشانی اینترنتی، پیوندها) قرار داده‌اند. در این پژوهش، تلاش شد تا به بررسی بهبود ربط با استفاده از مقایسه کارکرد جست‌وجو در عنوان و نشانی در امکانات موتور کاوش گوگل، پرداخته شود و نتایج پژوهش نشان داد که جست‌وجو بر اساس نشانی اینترنتی در گروه‌های موضوعی فنی مهندسی و جست‌وجو در عنوان در گروه‌های موضوعی اقتصاد نتایج مرتبط‌تری را به دنبال دارد و تفاوتی بین جست‌وجوی مبتنی بر عنوان و نشانی اینترنتی در گروه علوم پایه مشاهده نمی‌شود. در موارد هم‌خوانی نیز شاهد میزان ربط بالایی هستیم، به این ترتیب، هم‌خوانی کلیدواژه‌های موضوعی در عنوان و نشانی اینترنتی یک صفحه وب بیانگر ارتباط بیشتر آن صفحه با موضوع مورد نظر است و باعث بهبود ربط می‌شود.

تفاوت معناداری در بین تعداد کلیدواژه‌های موجود در عنوان و نشانی اینترنتی در بین گروه‌ها وجود دارد. بنابراین می‌توان نتیجه گرفت که تفاوت میزان ربط نتایج حاصل از جست‌وجو در عنوان و

نشانی اینترنتی بین گروه‌های موضوعی احتمالاً ریشه در ماهیت واژگان بین گروه‌ها و عدم وجود یک تعریف استاندارد برای مفاهیم و رسم‌الخط فارسی در وب است و همچنین در صورت استفاده از کلیدواژه‌های چند واژه‌ای مورد جست‌وجو میزان هم‌خوانی عنوان و نشانی اینترنتی کاهش می‌یابد و مدیران وب سایت‌ها نیز باید برای افزایش رتبه سایت خود به هم‌خوانی واژگانی بین عنوان و نشانی اینترنتی سایت، توجه بیشتری مبذول دارند. نکته دیگر استفاده از کلمات کلیدی به کار رفته به صورت کنترل شده و مرتبط در آن صفحه وب در آدرس آن صفحه وب و شاخه‌های موجود در نشانی اینترنتی است و دامنه‌هایی که نشان‌دهنده یک محدوده جغرافیایی هستند (مثل ir) می‌توانند در بهینه‌سازی نشانی اینترنتی برای نسخه‌های محلی موتورهای کاوش مانند گوگل فارسی مفید باشند و باعث افزایش رتبه سایت در موتورهای کاوشی مانند گوگل و افزایش ربط را مشمول شود.

به طور کلی، یافته‌های این پژوهش در زمینه جست‌وجوی منابع مرتبط و مفید وبی به کاربران در گروه‌های موضوعی اقتصاد، علوم پایه وگروه فنی مهندسی در ابزارهای کاوش با واسط فارسی و در زمینه طراحی بهتر و مناسب‌تر صفحات وب به طراحان وب‌های فارسی‌زبان با توجه به میزان حضور کلیدواژه در مکان‌های مختلف، یاری می‌رساند و همچنین می‌تواند متخصصان رایانه و طراحان موتورهای کاوش را در تدوین الگوریتم‌های جست‌وجو برای موتورهای کاوش و نیز صفحه رابط جست‌وجو که به شکل پیش‌فرض در اختیار کاربر قرار داده می‌شود یاری رساند و به شیوه‌های بهتر نمایه‌سازی موتورهای کاوش به منظور بازیابی مؤثر اطلاعات و نیز هدایت کاربر در بازیابی اطلاعات مرتبط‌تر، در کلیدواژه‌های موضوعی فارسی‌زبان در عنوان و نشانی اینترنتی کمک کند و تفاوت این پژوهش با سایر پژوهش‌های انجام‌شده در استفاده از زبان کلیدواژه‌های مورد جست‌وجو(که با واسط فارسی) در این پژوهش انجام گرفته است.

پیشنهادهای پژوهش:

۱. پیشنهادهایی به کاربران

در مجموع نتایج نشان می‌دهد که بسط جست‌وجو در حالت جست‌وجوی عنوانی و نشانی اینترنتی در تمام گروه‌های موضوعی مورد پژوهش می‌تواند موفقیت‌آمیز باشد.کاربران در وب می‌توانند با کنترل تعداد کلیدواژه‌های مورد جست‌وجو و با محدود کردن عبارات و کلیدواژه‌های مورد نیاز خود به جست‌وجو در عنوان (به خصوص در موارد مفهومی) و در موارد نیاز به اسامی خاص و تخصصی به طریق جست‌وجو در نشانی اینترنتی به نتایج بسیار دقیق‌تر و مرتبط‌تری در موتور کاوش گوگل دست یابند.

۲. به طراحان وب سایت‌ها و موتورهای کاوش

یافته‌های این پژوهش نشان داد که میان نتایج جست‌وجوهایی که کلیدواژه مورد جست‌وجو در عنوان و نشانی اینترنتی هم‌خوانی بیشتری وجود دارد، میزان ربط بالاتر است. هنگامی که واژه‌های مرتبط در هر دو منطقه عنوان و نشانی حضورداشته باشند، سایت مورد نظر از رتبه بالایی در بازیابی برخوردار می‌شود و نیز مشخص شد که وب‌سایت‌هایی که دارای کلیدواژه‌های تک واژه‌ای هستند از هم‌خوانی بالایی برخوردارند. بنابراین، طراحان صفحات وب را می‌توان تشویق نمود تا برای نام‌گذاری عنوان صفحات و نیز برای تدوین نشانی اینترنتی از کلیدواژه‌های کوتاه و ساده و معنادار استفاده نمایند. و موتورهای کاوش می‌توانند وزن‌دهی بیشتری را برای کلیدواژه‌های موجود در نشانی اینترنتی صفحات وب در نظر گیرند. همچنین، اگر به جای توجه زیاد به کلیدواژه‌های موجود در عنوان صفحات وب ، به کلیدواژه‌های موجود در نشانی اینترنتی نیز توجه نمایند، می‌توانند نتایج مرتبط‌تری را رتبه‌بندی کرده و در اختیار کاربر قرار دهند.

۳. برای پژوهش‌های آینده:

در این پژوهش فقط از دو گزینه جست‌وجو به روش "عنوان" و "نشانی اینترنتی" استفاده شد ، انجام پژوهشی در رابطه با مقایسه ربط از طریق جست‌وجو در دیگر مکان‌هایی خاص از جمله متن، نشانی اینترنتی و پیوندهای موجود در وب سایت‌ها که برای بازیابی اطلاعات مؤثرند ضروری به نظر می‌رسد و به منظور دسترسی بیشتر به اطلاعات مرتبط در گروه‌های موضوعی مختلف، لازم است مقایسه‌ای بین میزان ربط نتایج یافته شده در موتورهای جست‌وجو و نتایج یافته شده در دیگر ابزارهای بازیابی اطلاعات در وب مانند راهنماهای موضوعی با توجه به گروه موضوعی کاربران صورت گیرد و به لحاظ تعداد موتورهای کاوش در وب و استفاده آن‌ها توسط کاربران نامحدود، لازم است چنین پژوهشی در راستای اندازه‌گیری میزان ربط نتایج در چند موتور کاوش با رویکرد کار برمدار انجام پذیرد تا تواناترین موتور کاوش در ارائه نتایج مرتبط و عوامل مؤثر بر آن شناسایی شود.

جامعه آماری این پژوهش را کارشناسان استانداری همدان تشکیل می‌دهند، پیشنهاد می‌شود پژوهشی مشابه در سایر ارگان‌ها و مراکز علمی انجام شود و نتایج آن با نتایج این پژوهش مقایسه شود و همچنین برای تضمین کیفیت ارتباط و انتقال مؤثر پیام و حفظ سلامت زبان و رعایت استانداردهای آن یک استاندارد برای مفاهیم و رسم‌الخط فارسی در وب تعریف گردد.

کتابنامه:

علامی، پریسا و فتاحی رحمت اله(۱۳۹۱) تأثیر عنوان و نشانی اینترنتی در میزان ربط نتایج حاصل از بازیابی اطلاعات در موتورهای جستجو در دو حوزه علوم کشاورزی و علوم انسانی . علوم و فناوری اطلاعات. پژوهشنامه پردازش و مدیریت. ۲۸(۱) ۲۰۳–۲۲۴

فتاحی، رحمت الله(۱۳۸۵). " شناسایی و تحلیل واژگان عمومی در منابع وب: رویکردی نو به بسط عبارت جست‌وجو با استفاده از زبان طبیعی در موتورهای کاوش." *مطالعات تربیتی و روا نشناسی*، ۷(۱)۳۱–۵۲.

Allami P, Fattahi R. Comparing the Influence of Title and URL in Information Retrieval Relevance in Search Engines Results between Human Science and Agriculture Science. *Journal of Information Processing and Management*. 2012; 28 (1) :203-224.

Fattahi, Rahmatollah (2008). An alternative appproach to natural language query expansion in search engines: text analysis of non-topical terms in web documents. *Psychology Studies in Educationand*. 2008;7(1):31-52

Fattahi, Rahmatollah (2008). An alternative appproach to natural language query expansion in search engines. Concepción S. Wilson& Fletcher Cole Information Processing and Management 44 (2008) 1503–1516 .

Greg, R. Review of HotBot (Inktomi). ۲۰۰۴.

http://www.searchengineshowdown.com/features/hotbot/ (accessed۹ Feb ۲۰۰۸).

Kan, min-yen;

Kang, In-Ho, and Gil Chang Kim. ۲۰۰۴.Integration of multiple evidences based on a query type for web search. *Information Processing & Management* ۴۰, no. ۳ (may ۲۰۰۴): ۴۵۹-۴۷۸ .

Lawrence, Steve …[et al].۲۰۰۱. Persistence of Web References in Scientific Research. Computer magazine, ۲۰۰۱: v.۳۴, no.۲ ,۲۶-۳۱ .

SEO committee.(2008). Importance of Keywords in URLs and Search Engine Ranking. *Search engine journal*. Retrieved March 21, 2008, from http://www.searchenginejournal.com/importance-of-keywords-in-urls-and-search-engine-ranking/1019/

Shin, L- K, and D- R Karger.۲۰۰۴ Using urls and table layout for web classification tasks . *Proceedings of the ۱۳th international conference on World Wide Web* ۲۰۰۴.

Web position . (2008). Retrieved January 2, 2008 from http://en.kioskea.net/web/referencement.php3.

Yoshikiyo, Kato, and Kurohashi Sadao. (2008). " Classifying information sender of web documents." *Internet Research* ,18 (2) 191-203 .